2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR) | 979-8-3503-1184-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/MSR59073.2023.00055

# Insights into Female Contributions in Open-Source Projects

Arifa I. Champa Md Fazle Rabbi Minhaz F. Zibran
*Department of Computer Science, Idaho State University, USA*
{arifaislamchampa, mdfazlerabbi, minhazzibran}@isu.edu

Md Rakibul Islam
*University of Wisconsin - Eau Claire, USA*
islamm@uwec.edu

***Abstract*—This paper presents a large quantitative study of the contributions of females compared to males in open-source projects. Female participation is found substantially low and females are found more engaged in non-coding work compared to men. The findings are statistically significant and are derived from an in-depth analysis of over 10 thousand developers' contributions to more than 81 million different projects in the World of Code (WoC) infrastructure. The insights from this study are useful in addressing gender disparity in the field.**

***Index Terms*—gender, female, diversity, open-source, study**

## I. Introduction

Despite the early days' programmers being women, the lady Augusta Ada being recognized as the first programmer, the participation of women in computing has remained low compared to men. The same scenario exists in the fields of software development and software engineering. According to the US Bureau of Labor Statistics, women's employment in software development and programming roles have decreased from 2017 to 2019 and women make up only 18.7% of all software developers in the US in 2019 [1].

While women's participation in the software industry looks slightly better in first-world countries such as the US, the scenario appears much worse when we look at the global picture across countries. According to a report [2] in 2002, women made up only 1.1% of the contributors in FLOSS (Free/Libre and Open Source Software) projects. 15 years later, in 2017, GitHub surveyed over 5,500 open-source developers and users where only 5% of the contributors were identified as females [3]. This report further emphasized that gender diversity in the software engineering field is even worse than in the overall tech industry in general.

It is a general understanding that diverse teams outperform their homogeneous counterparts in a variety of ways with higher productivity, innovation, and overall success [4]–[7]. Hence, there is a rising interest among practitioners and researchers to examine gender diversity and gender bias in software development projects [7]–[11]. The ultimate goal behind all these recent efforts is to increase women's participation in the fields of software engineering and software development.

Towards this goal, we first need a clear understanding of the women's current contributions to software development, especially, the kinds of technologies and development tasks they are and aren't engaged in. These are what we investigate in this study. In particular, we address the following four research questions:

**RQ1:** *How prevalent are female contributors in popular open-source projects?*
— As described before, women's participation in the fields of software engineering and software development was reported very low in the past. We want to verify if and to what extent the scenario has changed over time, to gain an understanding of how much the fields have advanced in terms of gender diversity.

**RQ2:** *How do women's depth of engagement/contributions in open-source projects compare with those of men?*
— We estimate the depth of contributions/engagement of a contributor in terms of their dealing with coding and non-coding tasks. Comparing male and female contributors' involvement in coding and non-coding tasks will give us an insights into the depth of technical challenges these contributors take on.

**RQ3:** *Are there significant differences in the types of tasks (e.g., implementation of new features vs. fixing bugs in existing code) the male and female contributors are mostly engaged in?*
— Substantial disparities, if found in the categories of tasks the male and female contributors are engaged in, will indicate opportunities for capturing gender-wise interests and/or deficiencies in tackling different categories of tasks.

**RQ4:** *In which programming languages do the female developers contribute the most? How do their contributions in different languages change over time?*
— Understanding the female expertise and trends can inform the development of resources and training programs that cater to the specific needs and interests of female developers.

To address the aforementioned research questions, we analyze over 10 thousand contributors' nearly 21 million commits to more than 81 million different projects in the World of Code (WoC) infrastructure [12].

## II. Data Collection

We use WoC [12] as our primary source of data. WoC includes a curated dataset consisting of 173 million git repositories. One of WoC's primary areas of curation includes parsing file content to identify dependencies in 17 different programming languages. These 17 languages are listed in the second column from the left in Table I.

2574-3864/23/$31.00 ©2023 IEEE
DOI 10.1109/MSR59073.2023.00055

TABLE I
17 POPULAR PROJECTS FORMING OUR INITIAL DATASET

| Open-Source Project | Language | # of stars (in thousands) | Size (LOC) |
|---|---|---|---|
| gitlabhq_gitlabhq | Ruby | 23.1 | 5,460,688 |
| goldbergyoni_nodebestpractices | JavaScript | 85.4 | 32,090 |
| Seldaek_monolog | PHP | 20.1 | 17,743 |
| huihut_interview | C++ | 27.2 | 14,837 |
| Genymobile_scrcpy | C | 75.8 | 22,991 |
| avelino_awesome-go | Go | 95.2 | 4,620 |
| Snailclimb_JavaGuide | Java | 130 | 40,239 |
| donnemartin_system-design-primer | Python | 209 | 10,382 |
| yahoo_CMAK | Scala | 11.2 | 20,823 |
| qinwf_awesome-R | R | 5.3 | 798 |
| alacritty_alacritty | Rust | 44.2 | 28,226 |
| jscl-project_jscl | Lisp | 0.81 | 19,251 |
| TrinityCore_TDB_4.3.4_NLU | SQL | 0.062 | 1,110,955 |
| wrf-model_WRF | Fortran | 0.9 | 1,361,726 |
| JustArchiNET_ArchiSteamFarm | C# | 9 | 58,031 |
| NvChad_NvChad | Lua | 14 | 1,863 |
| shadowsocks_ShadowsocksX-NG | Swift | 30.9 | 38,369 |

The WoC dataset occupies over 250TB of disk space. For this study, we need a manageable subset. Thus, for each of the 17 programming languages that WoC identifies, we select the most popular (based on star-rating on GitHub) project (written *primarily* in that language), which is available in both GitHub and WoC. These 17 popular projects constitute our initial dataset to address research questions RQ1 and RQ2. Table I presents these 17 projects' sizes (in number of lines of code) and their star-rating popularity on GitHub as of January 2023. This initial dataset is further extended systematically as we continue our analyses to address the research questions RQ3 and RQ4. We end up analyzing 20,997,331 commits made by 10,732 different developers across 81,722,661 distinct projects.

### A. Author/Contributor and Gender Identification

To identify the authors/developers/contributors of a project, we use the project to author mapping in WoC. A total of 10,898 distinct authors/contributors are identified to have contributed to the 17 projects listed in Table I. To detect the gender of an author, we use Wiki-Gendersort [13]. Among the 10,898 contributors, 10,255 are identified as males, 477 are as females, and the rest 166 are classified as unisex or unknown.

### B. Change Characterization

Software projects include source code written in files, which we refer to as code files (CF). The language in which a CF is written is determined based on the the file extensions. For example, a file having a `.c` extension is considered a CF written in C. Files having a `.md` extension are considered non-code files (NF). A typical commit may make changes to CF or/and NF. For this study, we keep account of the number of CF and NF being affected by a certain commit. For convenience, let $\delta$ and $\gamma$ respectively denote the number of *CF* and *NF* affected/changed by commits made per author/contributor on average.

## III. ANALYSIS AND FINDINGS

The procedural steps for data collection and analyses in our study are summarized in Figure 1. The steps in phase 1 through phase 4 respectively are associated with addressing research questions RQ1 through RQ4.

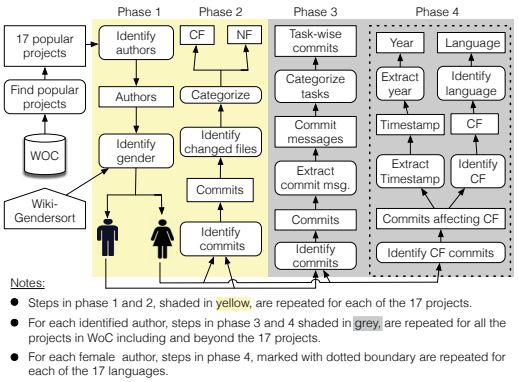

Fig. 1. Procedural steps of our study

TABLE II
MALE AND FEMALE CONTRIBUTIONS IN POPULAR PROJECTS

| Prog. Lang. | Male authors/contributors | | | | Female authors/contributors | | | |
|---|---|---|---|---|---|---|---|---|
| | | Per author avg # of | | | | Per author avg # of | | |
| | | | Changed | | | | Changed | |
| | % | commits | NF ($\gamma$) | CF ($\delta$) | % | commits | NF ($\gamma$) | CF($\delta$) |
| Ruby | 93.49 | 38.01 | 58.39 | 390.55 | 5.14 | 53.20 | 58.15 | 360.52 |
| JavaScr. | 94.79 | 11.92 | 56.00 | 69.15 | 3.29 | 30.67 | 196.08 | 0 |
| PHP | 95.57 | 5.05 | 0.79 | 17.40 | 2.22 | 2.06 | 0.56 | 12.78 |
| C++ | 92.50 | 9.03 | 11.43 | 1.70 | 7.50 | 15.67 | 16.67 | 3.33 |
| C | 94.54 | 14.81 | 4.63 | 20.53 | 3.83 | 2.57 | 0.57 | 4.00 |
| Go | 96.49 | 2.81 | 2.80 | 0.07 | 2.53 | 2.16 | 2.13 | 0 |
| Java | 94.20 | 7.67 | 34.50 | 0.88 | 4.22 | 31.13 | 57.25 | 0.56 |
| Python | 95.56 | 4.26 | 6.64 | 1.20 | 2.96 | 2.25 | 1.83 | 0.25 |
| Scala | 91.26 | 5.47 | 0.51 | 21.69 | 7.10 | 4.62 | 0.38 | 15.00 |
| R | 93.55 | 4.95 | 4.71 | 0.21 | 5.16 | 2.50 | 2.50 | 0 |
| Rust | 95.64 | 11.02 | 3.54 | 22.73 | 2.25 | 6.19 | 2.19 | 14.19 |
| Lisp | 96.88 | 45.48 | 0.89 | 81.42 | 3.13 | 24.50 | 0 | 82.50 |
| SQL | 85.54 | 13.49 | 0.03 | 159.15 | 1.20 | 3.00 | 0 | 3.00 |
| Fortran | 76.90 | 48.21 | 0 | 98.34 | 21.41 | 13.83 | 0.30 | 63.34 |
| C# | 93.67 | 147.49 | 5.24 | 120.00 | 3.80 | 1.33 | 0 | 2.33 |
| Lua | 98.06 | 13.66 | 3.05 | 28.48 | 1.94 | 1.50 | 0 | 2.50 |
| Swift | 93.68 | 11.04 | 4.75 | 79.89 | 3.16 | 4.33 | 2.00 | 65.67 |
| **Overall** | 93.08 | 23.20 | 11.64 | 65.49 | 4.76 | 11.85 | 20.04 | 37.06 |

For each of the 17 projects (identified in Table I), we compute the average number of commits they collectively made to the project, and the average number of CF and NF affected by commits per author (i.e., $\delta$ and $\gamma$, respectively) as presented in Table II. The same set of information for the female authors/contributors are also included in the table.

### A. Prevalence of Male or Female Contributors (RQ1)

A total of 10,898 distinct authors/contributors are identified to have contributed to the 17 projects listed in Table I. Among the 10,898 contributors, 10,255 are identified as males, 477 are as females, and the rest 166 are classified as unisex or unknown. This 1.57% of contributors classified as unisex or unknown are excluded from our study. None of the 10,898 authors is found to have contributed to multiple of the 17 projects.

Clearly the total number of distinct male contributors (10,898) is much higher than the total number of female contributors (477). For each of the 17 projects (except for one) the proportion of female contributors is only 7.5% or

lower. The only exception is the Fortran project, where females constitute 21.41% of the contributors.

Fortran is arguably the first high level programming language and it is interesting that the female developers' proportion in the Fortran project (i.e., wrf-model_WRF) is much higher compared to their participation in projects primarily developed in other programming languages. This is among the reasons that inspires us pursue the research question RQ4, which we address later in Section III-D.

Now, if we consider per person number of commits on average, male contributors appear to have higher number of commits. For most of the projects (except for four) the average per person number of commits made by the male contributors is consistently higher than those made by females. The exceptions are the four projects primarily written in Ruby, JavaScript, C++, and Java, where the per person number of commits made by females are found higher than those made by the male contributors. Overall, across all the projects, the per person average number of commits by male contributors is 23.20, which is close to double the contributions of females (per person 11.85 commits only). A Mann-Whitney-Wilcoxon (MWW) test [14] over the distributions of per person average number of commits by male and female contributors confirms the statistical significance ($p = 0.03, \alpha = 0.05$) of our observation. We, therefore, derive the answer to the research question RQ1 as follows:

**Ans. to RQ1:** *Female participation in open-source projects is generally low. The proportion of female contributors is only 7.5% or lower. Females make fewer commits to the projects compared to male contributors.*

### B. Technical Depth of Contributions (RQ2)

We characterize the technical depth/engagement of a commit by accounting the number of CF and NF affected/changed by that commit. A contributor making commits that affect more CF than NF is considered to have made more technical/coding contributions than non-coding contributions.

For each of the 17 projects (listed in Table I), all commits made by male and female contributors are selected separately. Then, the files that are modified through these commits are identified and classified into CF or NF according to the characterization described in Section II-B.

For the 17 projects, if we look at $\delta$ (i.e., per author average number of affected CF) and $\gamma$ (i.e., per author average number of affected NF) as presented in Table II, we note the following observations: (a) for almost all the projects (with two exceptions for C++ and Lisp projects), $\delta$ is substantially higher for male contributors compared to females. Across all projects, overall $\delta$ for male contributors is 65.49, while for females $\delta$ is only 37.06. (b) For almost all the projects (with four exceptions for JavaScript, C++, Java and Fortran projects), $\gamma$ is also higher for male contributors compared to females. However, across all projects, overall $\gamma$ for females is 20.04, which is higher than that for males (11.64). (c) When we focus on male contributors only, in most (11) of the projects $\delta$ is substantially higher than $\gamma$, which is also reflected on the overall values across all the projects ($\delta = 65.49$ while $\gamma = 11.64$ only). (d) A similar scenario is also visible for female contributors but for females, the difference between overall $\delta$ (37.06) and $\gamma$ (20.04) across all the projects is much smaller compared to males. Based on these observations, we formulate the answer to research question RQ2 as follows:

**Ans. to RQ2:** *Females make less contributions in both coding and non-coding tasks compared to males. The ratio of coding to non-coding contributions is much smaller for females compared to male contributors.*

### C. Task Assignment (RQ3)

To address RQ3, we examine the contributions of both male and female authors across five different categories of tasks, which are: 'bug fixing', 'energy aware', 'new feature', 'refactoring', and 'security-related' tasks. We identify 0.57 million commits made by female authors and almost 21 million commits made by male authors in *all* the projects of WoC (including the 17 projects mentioned before).

We then categorize the commits by classifying corresponding commit messages in the five task categories using a keyword based substring search method adopted from the work of Islam and Zibran [15]–[18]. Thus we deal with 565,239 commits by 477 female authors to 2,633,275 distinct projects and 20,997,331 commits by 10,255 male authors to 79,089,386 distinct projects. Among them 351,644 female commits and 14,356,162 male commits are discarded because those do not fit into any of the five task categories.

TABLE III
MALE AND FEMALE COMMITS FOR DIFFERENT TASK CATEGORIES

| Categories of Tasks | Number of commits | | | |
|---|---|---|---|---|
| | By male developers | | By female developers | |
| | Total | Average | Total | Average |
| Bug fixing | 2,271,537 | 221.51 | 80586 | 168.94 |
| Energy aware | 4504 | 0.44 | 63 | 0.13 |
| New feature | 2,894,381 | 282.24 | 106323 | 222.9 |
| Refactoring | 157,081 | 15.32 | 5091 | 10.67 |
| Security-related | 1,313,666 | 128.1 | 21512 | 45.1 |

Table III presents the number of commits and the average number of male and female commits per author for the five task categories. As seen in the table, for each category of tasks, per person average number of commits is consistently higher for male contributors compared to females. Especially, for the 'security-related' tasks, male authors are found to have made nearly three times the commits made by the female authors (i.e., 128.1 vs. 45.1). Pairwise MWW tests ($\alpha = 0.05$) between the male and female contributions/commits performed separately for each of the five categories of tasks confirm statistical significance of our observation with $p = 0.00001$. We, therefore, derive the answer to RQ3 as follows:

**Ans. to RQ3:** *Female contributors consistently make fewer commits than males across all task categories, with a very large disparity in contributions to 'security-related' tasks.*

TABLE IV
CONTRIBUTION OF 477 FEMALES IN 17 PROGRAMMING LANGUAGES

| Language | # of commits | First commit | Language | # of commits | First commit |
|---|---|---|---|---|---|
| Ruby | 576623 | 2008 | R | 27742 | 2011 |
| JavaScript | 411048 | 2006 | Rust | 12863 | 2008 |
| PHP | 378286 | 2006 | Lisp | 12756 | 2009 |
| C++ | 361176 | 1992 | SQL | 10860 | 2006 |
| C | 305761 | 1992 | Fortran | 9253 | 2008 |
| Go | 202124 | 2012 | C# | 3790 | 2012 |
| Java | 64906 | 2010 | Lua | 3047 | 2012 |
| Python | 44834 | 2006 | Swift | 874 | 2015 |
| Scala | 38488 | 2011 | **Total** | **2,464,431 commits** | |

### D. Programming Languages Used by Females (RQ4)

For each of the 477 female developers identified before, we capture the commits affecting CF in all projects in WoC including and beyond the 17 projects discussed before. Thus, we obtain 2,464,431 CF commits across 2,633,275 projects. For each of these commits, we extract year from timestamp and for each of the affected CF, we identify the language the CF is written in.

Table IV presents the number of commits those female contributors made resulting in changes to CF written in the 17 languages. As we see, the female commits made changes to CF written in all 17 languages, which include modern languages (e.g., Go, Swift, Rust) as well as older languages (e.g., Fortran, C). Table IV also presents the year in which code written in each language was first affected by a female commit and we observe interesting variations.

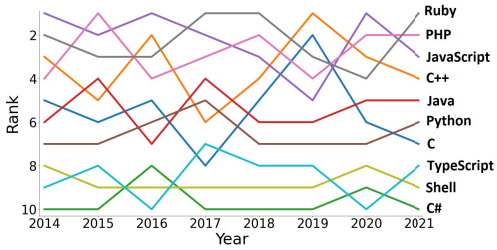

Fig. 2. Trends of the top ten languages females contribute to

GitHub recently published a trend of top ten programming languages in which contributions to GitHub projects were made the most [19]. In Figure 2, we present a similar trend among the 477 WoC female contributors in our study. Again, unlike the GitHub trend that includes developers of all genders, the female contributors in our study exhibit substantial variations in the languages they contribute most in different years. Based on the observation, we derive the answer to RQ4 as follows:

**Ans. to RQ4:** *Females contribute in all programming languages. Over long period, Ruby, JavaScript, PHP, C++, C, and Go are the top six languages the females contribute. However, their top contributing languages vary across years.*

## IV. THREATS TO VALIDITY

In estimating contributions, we relied on the number of commits and the number of affected CF and NF without taking into account the dispersion and sizes of changes made to the files. We plan to address this limitation in the future.

Due to some difficulties in extracting individual developers' NamSor [20] gender mappings in WoC, we opted in using Wiki-Gendersort. This algorithm is reported to have 97.07% accuracy in identifying genders in NamSor names [13]. The exclusion of 1.57% of contributors (Section III-A) classified as unisex or unknown should not have a significant impact on our results.

## V. RELATED WORK

There have been many studies on open-source projects [21]–[25] but only a few focused on female participation and contribution. Surveys over years reported low participation/contribution of women in open-source projects [2], [3], [26], [27] while some reports also indicated that female contributors primarily engaged in non-coding tasks [26], [27].

Based on a study on code review repositories of 10 popular open source projects, Bosu and Sultana [28] identified less than 10% female developers and even fewer female core developers. A recent survey of 51 articles reported that women are less likely to be core developers and often make non-code contributions [29]. Canedo et al. [30] studied core developers of 711 open source projects and reported that only 2.3% core developers were women while only 45 out of 711 systems had at least one woman core developer.

Our work is unique from the aforementioned ones in several ways. Ours is not a survey. Compared to the earlier studies, ours is a substantially large scale work. Our first research question was investigated in the past but not at the scale of our study and the rest three research questions were never explored before. Similar to earlier work, we also found very low participation of females with only 7.5% contributors being women. However, we find that women contribute to both coding and non-coding tasks, although both kinds of contributions still remain lower than male contributions.

## VI. CONCLUSION

In this paper, we have presented a large quantitative study of the prevalence, engagement, and expertise of female contributors (compared to males) in open-source projects in WoC [12].

We find that the proportion of female contributors is only 7.5% or lower. Females make fewer commits compared to males irrespective of categories of tasks. Females make less contributions in both coding and non-coding tasks compared to males. Contributions in coding compared to non-coding tasks are higher for both males and females, but the difference is comparatively much smaller for females. Female developers are found to have contributed in all the programming languages with the most contributions in Ruby, JavaScript, PHP, C++, C, and Go.

The results are derived from an in-depth analysis of over 10 thousand developers' nearly 21 million commits to more than 81 million different projects. The results are verified in the light of statistical significance. In future, we plan to address the limitations identified in Section IV and extend this work with a qualitative investigation of the reasons behind the findings of this study.

## REFERENCES


[1] "An analysis of women in computing," Grand Canyon University, https://www.gcu.edu/blog/gcu-experience/analysis-women-computing, 2020, (Verified Jan 31, 2023).

[2] R. A. Ghosh, R. Glott, B. Krieger, and G. Robles, "Free/libre and open source software: Survey and study," pp. 1–68, 2002.

[3] K. Finley, "Diversity in open source is even worse than in tech overall," *Wired Magazine Website, https://www.wired.com/2017/06/diversity-open-source-even-worse-tech-overall*, 2017.

[4] C. R. Østergaard, B. Timmermans, and K. Kristinsson, "Does a different view create something new? the effect of employee diversity on innovation," *Research Policy*, vol. 40, no. 3, pp. 500–509, 2011.

[5] P. C. Earley and E. Mosakowski, "Creating hybrid team cultures: An empirical test of transnational team functioning," *Academy of Management Journal*, vol. 43, no. 1, pp. 26–49, 2000.

[6] P. Tourani, B. Adams, and A. Serebrenik, "Code of conduct in open source projects," in *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 2017, pp. 24–33.

[7] B. Vasilescu, D. Posnett, B. Ray, M. G. van den Brand, A. Serebrenik, P. Devanbu, and V. Filkov, "Gender and tenure diversity in github teams," in *33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 3789–3798.

[8] N. Robson, "Diversity and decorum in open source communities," in *2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2018, pp. 986–987.

[9] G. Catolino, F. Palomba, D. A. Tamburri, A. Serebrenik, and F. Ferrucci, "Gender diversity and women in software teams: How do they affect community smells?" in *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, 2019, pp. 11–20.

[10] N. Imtiaz, J. Middleton, J. Chakraborty, N. Robson, G. Bai, and E. Murphy-Hill, "Investigating the effects of gender bias on github," in *2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)*, 2019, pp. 700–711.

[11] Y. Wang and D. Redmiles, "Implicit gender biases in professional software development: An empirical study," in *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, 2019, pp. 1–10.

[12] A. Mockus, A. Nolte, and J. Herbsleb, "MSR Mining Challenge: World of Code," in *Proceedings of the International Conference on Mining Software Repositories (MSR 2023)*, 2023.

[13] N. Bérubé, G. Ghiasi, M. Sainte-Marie, and V. Larivière, "Wikigendersort: Automatic gender detection using first names in wikipedia," 2020.

[14] F. Ramsey and D. Schafer, *The statistical sleuth: a course in methods of data analysis*. Cengage Learning, 2012.

[15] M. R. Islam and M. F. Zibran, "Exploration and exploitation of developers' sentimental variations in software engineering," *International Journal of Software Innovation (IJSI)*, vol. 4, no. 4, pp. 35–55, 2016.

[16] M. R. Islam and M. F. Zibran, "Towards understanding and exploiting developers' emotional variations in software engineering," in *14th IEEE International Conference on Software Engineering Research, Management and Applications (SERA)*, 2016, pp. 185–192.

[17] M. R. Islam and M. F. Zibran, "What changes in where? an empirical study of bug-fixing change patterns," *ACM SIGAPP Applied Computing Review*, vol. 20, no. 4, pp. 18–34, 2021.

[18] M. R. Islam and M. F. Zibran, "How bugs are fixed: Exposing bug-fix patterns with edits and nesting levels," in *Proceedings of the 35th ACM Symposium on Applied Computing*, 2020, pp. 1523–1531.

[19] "The top programming languages in 2022: The state of the octoverse," https://octoverse.github.com/2022/top-programming-languages, (Verified: Jan 2023).

[20] "Namsor, name checker for gender, origin and ethnicity determination," https://namsor.app, (Verified Jan 2023).

[21] M. Islam and M. Zibran, "On the characteristics of buggy code clones: A code quality perspective," in *12th IEEE Intl. Workshop on Software Clones*, 2018, pp. 23 – 29.

[22] M. Islam and M. Zibran, "How bugs are fixed: Exposing bug-fix patterns with edits and nesting levels," in *35th ACM/SIGAPP Symposium on Applied Computing*, 2020, pp. 1523–1531.

[23] M. Islam and M. Zibran, "What changes in where? an empirical study of bug-fixing change patterns," *ACM Applied Computing Review*, vol. 20, no. 4, pp. 18–34, 2021.

[24] R. Joseph, M. Zibran, and F. Eishita., "Choosing the weapon: A comparative study of security analyzers for android applications," in *Intl. Conference on Software Engineering, Management and Applications*, 2021, pp. 51–57.

[25] A. Rajbhandari, M. Zibran, and F. Eishita, "Security versus performance bugs: How bugs are handled in the chromium project," in *Intl. Conference on Software Engineering, Management and Applications*, 2022, pp. 70–76.

[26] G. Robles, L. A. Reina, J. M. González-Barahona, and S. D. Domínguez, "Women in free/libre/open source software: The situation in the 2010s," in *IFIP International Conference on Open Source Systems*, 2016, pp. 163–173.

[27] A. Mani and R. Mukherjee, "A study of foss 2013 survey data using clustering techniques," in *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2016, pp. 118–121.

[28] A. Bosu and K. Z. Sultana, "Diversity and inclusion in open source software (oss) projects: Where do we stand?" in *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2019, pp. 1–11.

[29] B. Trinkenreich, I. Wiese, A. Sarma, M. Gerosa, and I. Steinmacher, "Women's participation in open source software: A survey of the literature," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 31, no. 4, pp. 1–37, 2022.

[30] E. D. Canedo, R. Bonifácio, M. V. Okimoto, A. Serebrenik, G. Pinto, and E. Monteiro, "Work practices and perceptions from women core developers in oss communities," in *Proceedings of the 14th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2020, pp. 1–11.